

**Towards accurate machine learning predictions of properties
of the P–O bond cleaving in ATP upon enzymatic hydrolysis**

Igor V. Polyakov, Kirill D. Miroshnichenko, Tatiana I. Mulashkina,
Alexander A. Moskovsky, Ekaterina I. Marchenko and Maria G. Khrenova

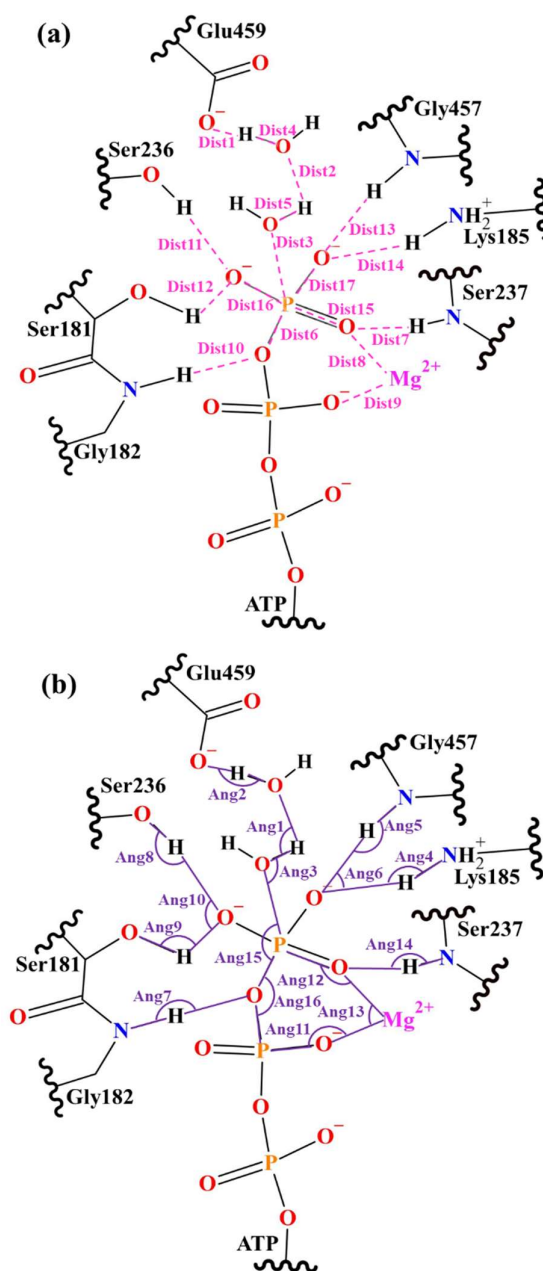


Figure S1. Geometry parameters, interatomic distances (a) and angles (b), selected as features for ML models.

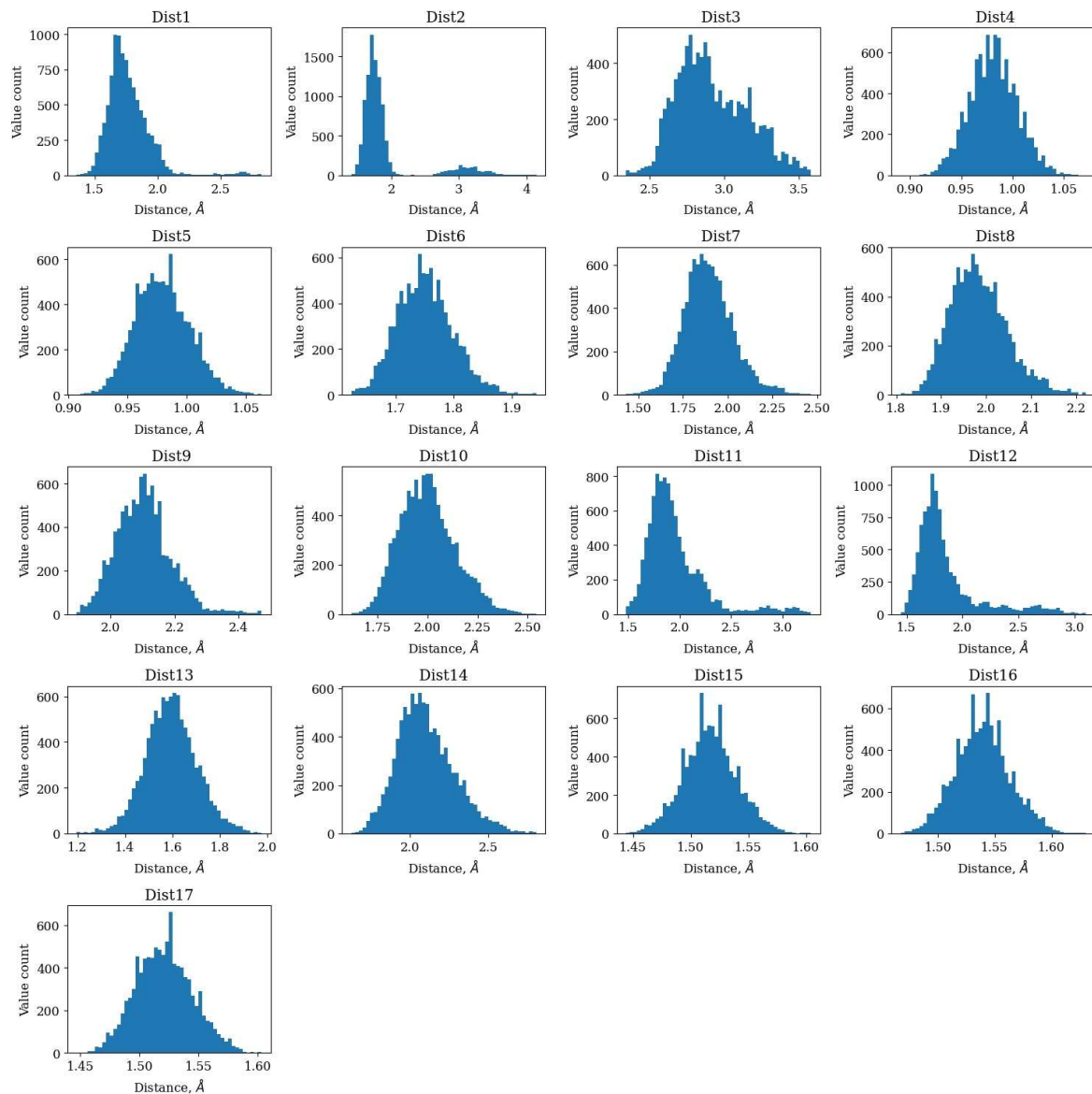


Figure S2. Distributions of interatomic distances selected as ML model features.

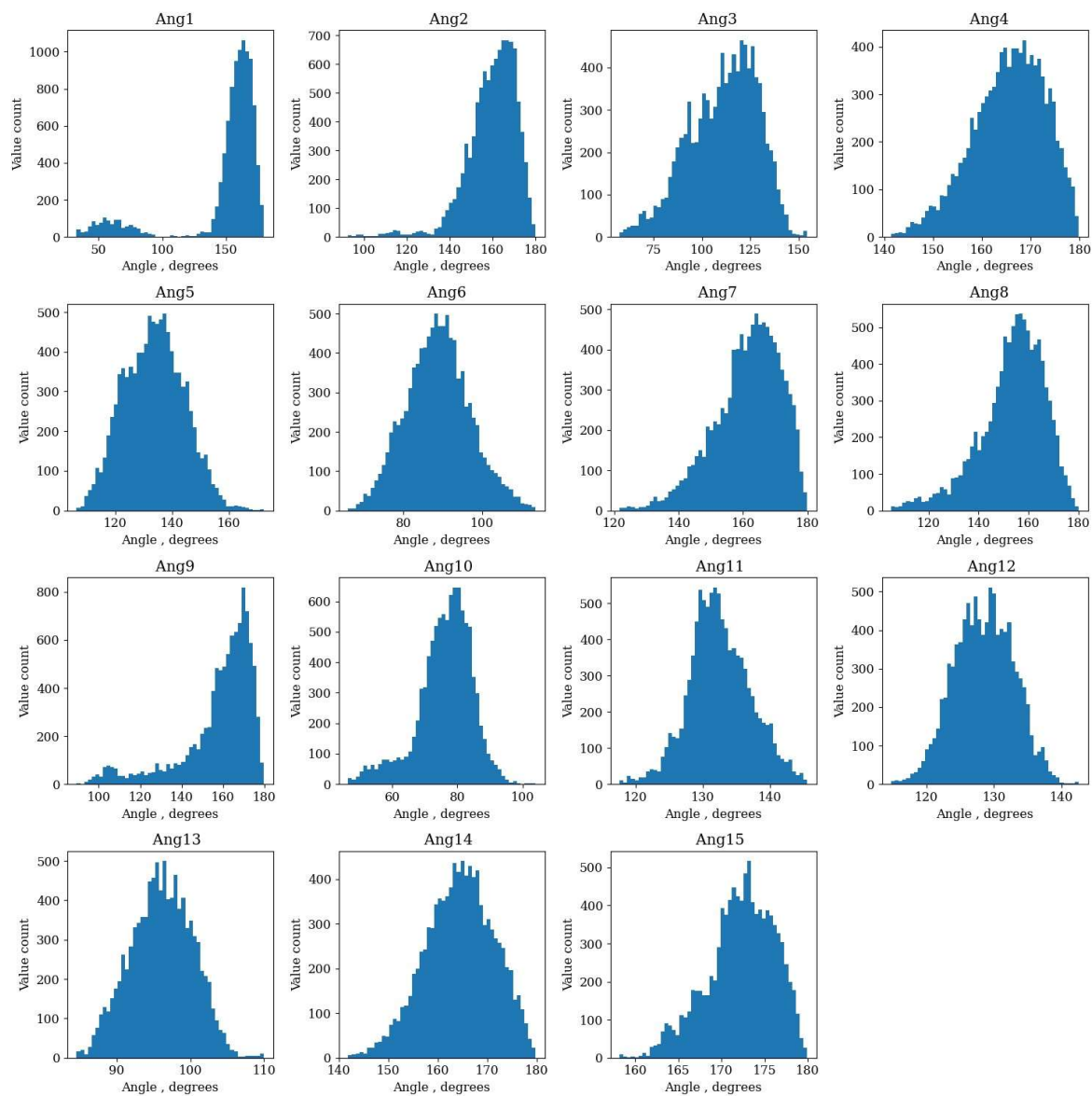


Figure S3. Distributions of angles selected as ML model features.

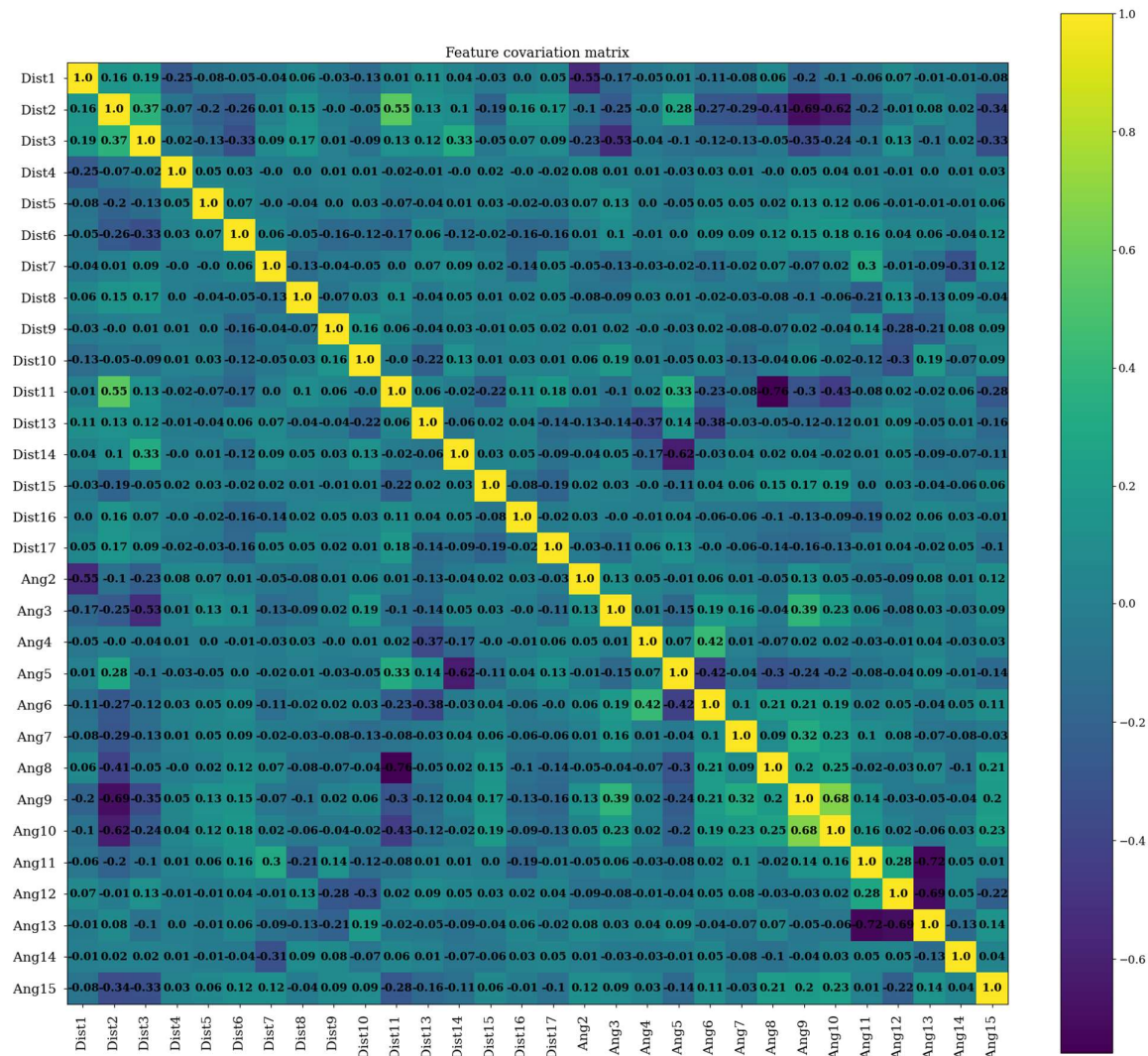


Figure S4. Feature correlation matrix.

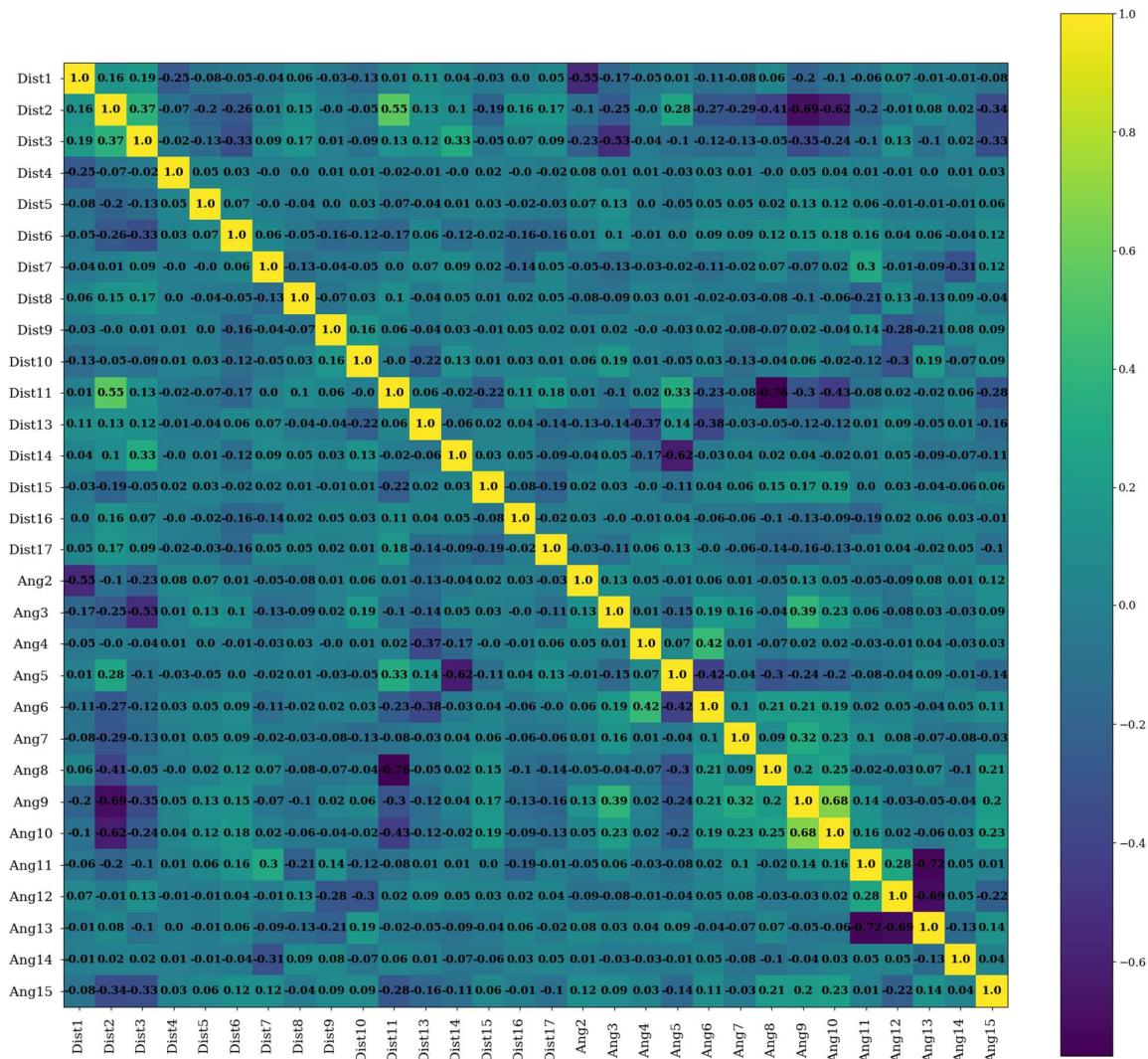


Figure S5. Feature correlation matrix with excluded Ang1 and Dist12.

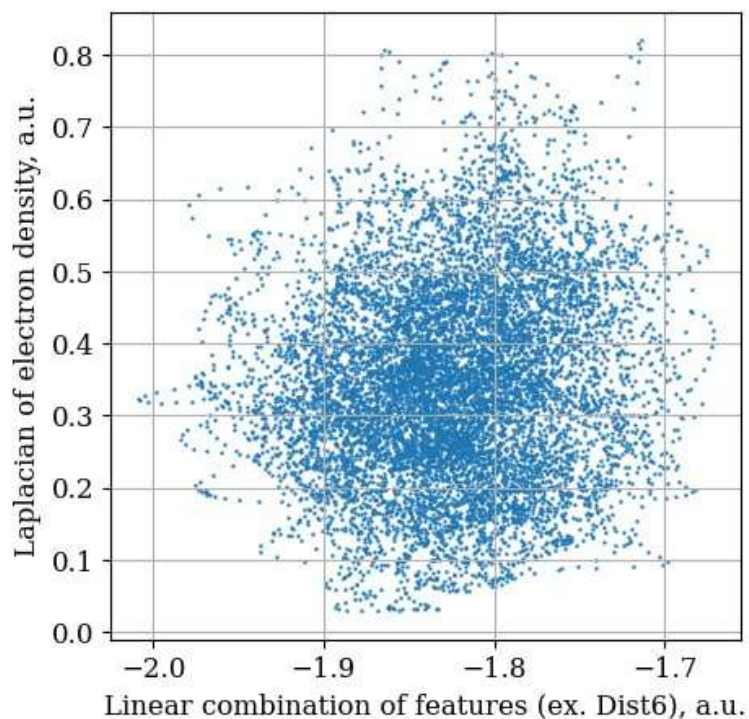


Figure S6. Laplacian of electron density calculated at the BCP of cleaving P_G–O_{3B} bond against the sum of all features excluding Dist6 (cleaving P_G–O_{3B} bond) calculated at 10000 QM/MM MD frames of the ES complex trajectory.

Table S1. Mean, maximum, minimum and standard deviation of feature values.

Feature	Mean $\pm$ st.dev., Å/°	Min, Å / deg.	Max, Å / deg.
Dist1	1.77 $\pm$ 0.17	1.36	2.83
Dist2	1.90 $\pm$ 0.48	1.41	4.14
Dist3	2.92 $\pm$ 0.24	2.34	3.58
Dist4	0.98 $\pm$ 0.02	0.89	1.07
Dist5	0.98 $\pm$ 0.02	0.91	1.06
Dist6	1.75 $\pm$ 0.05	1.62	1.94
Dist7	1.89 $\pm$ 0.14	1.43	2.46
Dist8	1.98 $\pm$ 0.06	1.81	2.22
Dist9	2.10 $\pm$ 0.09	1.89	2.47
Dist10	2.00 $\pm$ 0.14	1.62	2.54
Dist11	1.96 $\pm$ 0.30	1.48	3.27
Dist12	1.85 $\pm$ 0.29	1.45	3.10
Dist13	1.60 $\pm$ 0.11	1.20	1.98
Dist14	2.10 $\pm$ 0.18	1.62	2.81
Dist15	1.52 $\pm$ 0.02	1.44	1.60
Dist16	1.54 $\pm$ 0.02	1.47	1.63
Dist17	1.52 $\pm$ 0.02	1.45	1.60
Ang1	149.0 $\pm$ 33.7	32.8	179.5
Ang2	159.6 $\pm$ 11.7	92.7	179.9
Ang3	111.0 $\pm$ 17.7	57.4	154.4
Ang4	165.5 $\pm$ 7.4	141.4	179.8
Ang5	133.6 $\pm$ 10.6	106.3	172.3

Ang6	88.82±8.1	66.0	113.5
Ang7	161.3±10.3	121.6	179.9
Ang8	15.4±13.0	105.1	180.0
Ang9	157.8±17.9	89.3	179.8
Ang10	76.1±8.6	46.4	104.0
Ang11	132.5±4.6	117.6	145.5
Ang12	128.5±4.4	114.8	142.5
Ang13	96.1±4.2	84.6	110.0
Ang14	164.3±6.9	141.8	179.7
Ang15	172.2±3.9	158.2	179.9

Table S2. Regression coefficients and contribution

Feature	Coefficient	Contribution, %
Dist6	-0.132508	68.1
Ang13	-0.008479	4.4
Ang12	-0.007673	3.9
Dist3	-0.005861	3.0
Ang4	0.004091	2.1
Dist9	-0.004049	2.1
Dist13	0.003933	2.0
Dist15	-0.003781	1.9
Ang11	-0.003565	1.8
Ang15	-0.003501	1.8
Ang8	0.002143	1.1
Dist16	-0.001797	<1.0
Dist17	-0.001787	<1.0
Ang3	-0.001596	<1.0
Dist11	0.001406	<1.0
Ang7	0.001308	<1.0
Dist10	-0.001218	<1.0
Dist2	0.001072	<1.0
Dist14	0.000763	<1.0
Dist1	0.000735	<1.0
Dist7	0.000628	<1.0
Dist8	0.000557	<1.0
Ang6	0.000499	<1.0
Ang10	0.000437	<1.0
Ang14	-0.000407	<1.0
Dist4	0.000397	<1.0
Ang5	0.00182	<1.0
Ang9	-0.000073	<1.0
Ang2	0.000067	<1.0
Dist5	0.000043	<1.0

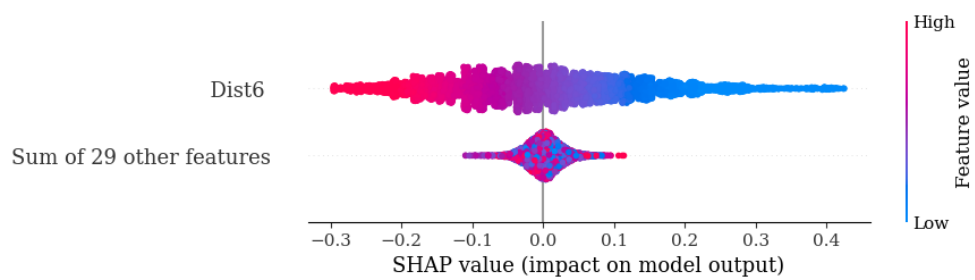


Figure S7. The beeswarm SHAP plot calculated for the LGBM model.